

SHORT COMMUNICATION

Limited evidence for adaptive evolution and functional effect of allelic variation at rs702424 in the promoter of the *TAS2R16* bitter taste receptor gene in Africa

Michael C Campbell¹, Alessia Ranciaro¹, Daniel Zinshteyn^{2,11}, Renata Rawlings-Goss¹, Jibril Hirbo¹, Simon Thompson¹, Dawit Woldemeskel^{1,3}, Alain Froment⁴, Sabah A Omar⁵, Jean-Marie Bodo⁶, Thomas Nyambo⁷, Gurja Belay³, Dennis Drayna⁸, Paul AS Breslin^{9,10} and Sarah A Tishkoff^{1,2}

Bitter taste perception, mediated by receptors encoded by the *TAS2R* loci, has important roles in human health and nutrition. Prior studies have demonstrated that nonsynonymous variation at site 516 in the coding exon of *TAS2R16*, a bitter taste receptor gene on chromosome 7, has been subject to positive selection and is strongly correlated with differences in sensitivity to salicin, a bitter anti-inflammatory compound, in human populations. However, a recent study suggested that the derived G-allele at rs702424 in the *TAS2R16* promoter has also been the target of recent selection and may have an additional effect on the levels of salicin bitter taste perception. Here, we examined alleles at rs702424 for signatures of selection using Extended Haplotype Homozygosity (EHH) and F_{ST} statistics in diverse populations from West Central, Central and East Africa. We also performed a genotype–phenotype analysis of salicin sensitivity in a subset of 135 individuals from East Africa. Based on our data, we did not find evidence for positive selection at rs702424 in African populations, suggesting that nucleotide position 516 is likely the site under selection at *TAS2R16*. Moreover, we did not detect a significant association between rs702424 alleles and salicin bitter taste recognition, implying that this site does not contribute to salicin phenotypic variance. Overall, this study of African diversity provides further information regarding the genetic architecture and evolutionary history of a biologically-relevant trait in humans.

Journal of Human Genetics (2014) 59, 349–352; doi:10.1038/jhg.2014.29; published online 1 May 2014

Bitter taste receptors, encoded by the *TAS2R* genes, are involved in a number of biological processes, including oral sensory perception^{1–3} and immune response,^{4–6} adding a level of complexity to our understanding of these receptors' function. Prior genetic analyses have reported strong signatures of selection at *TAS2R* loci on chromosome 7,^{7–11} suggesting these genes have been functionally important during human evolution. For example, the derived T-allele at site 516 within the coding exon of *TAS2R16* has been identified as a target of positive selection, and has been correlated with increased cell surface expression of the *TAS2R16* receptor and heightened sensitivity to salicin bitterness.^{8,10} A recent genetic study of global populations (including six African populations) also detected long-range linkage disequilibrium (LD) on chromosomes with both the derived T-allele at site 516 and the derived G-allele at rs702424 in the *TAS2R16* promoter, and inferred recent selection for haplotypes carrying these two mutations.⁹ To test for signatures of selection within the *TAS2R16*

promoter, we analyzed allelic variation at rs702424 together with sequence data from the *TAS2R16* coding exon in a large set of diverse populations from West Central, Central and East Africa. We also tested for an association between genetic variability at rs702424 and threshold levels of salicin taste recognition in individuals from East Africa.

We examined allelic variation (G/A) at rs702424 from the 1M Illumina Duo array genotyped in 697 individuals originating from 44 populations in West Central, Central and East Africa (Supplementary Table 1). African populations with shared genetic similarity, as well as cultural and/or linguistic properties (for example, Khoesan-speaking hunter-gatherers, Niger-Kordofanian speakers and Afroasiatic speakers) were pooled together for subsequent analyses as described by Campbell *et al.*⁸ and Tishkoff *et al.*¹². We found that the derived G-allele at rs702424 ranged in frequency from 4.2% in the Bulala population from Chad (Central Africa) to 32.5% in the Luo Nilo-

¹Department of Genetics, University of Pennsylvania, Philadelphia, PA, USA; ²Department of Biology, University of Pennsylvania, Philadelphia, PA, USA; ³Biology Department, Addis Ababa University, Addis Ababa, Ethiopia; ⁴UMR 208, IRD-MNHN, Museedel'Homme, Paris, France; ⁵Kenya Medical Research Institute, Centre for Biotechnology Research and Development, Nairobi, Kenya; ⁶Ministry of Scientific Research and Innovation, Yaounde, Cameroon; ⁷Department of Biochemistry, Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania; ⁸National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Rockville, MD, USA; ⁹Monell Chemical Senses Center, Philadelphia, PA, USA and ¹⁰Department of Nutritional Sciences, Rutgers University, New Brunswick, NJ, USA

¹¹Current address: Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, USA.

Correspondence: Professor SA Tishkoff, Department of Genetics, University of Pennsylvania, 415 Curie Boulevard, Philadelphia, PA 19104, USA.

E-mail: tishkoff@mail.med.upenn.edu

Received 28 January 2014; revised 3 April 2014; accepted 7 April 2014; published online 1 May 2014

Saharan speakers from Kenya (East Africa) (Supplementary Table 2). To investigate signals of selection based on allele frequency differences, we estimated average F_{ST} among populations at rs702424¹³ and compared this observed estimate with a genome-wide distribution of F_{ST} values derived from (a) 1069 randomly sampled single-nucleotide polymorphisms (SNPs) from the 1M Illumina Duo array and (b) HapMap Phase III SNP data in African populations. These analyses indicated that our observed F_{ST} estimate ($F_{ST}=0.01$) at rs702424 was not unusual compared with the empirical distribution (occurring at the 27th percentile for both the 1M Illumina and HapMap SNP data), implying the absence of extensive population divergence at this site. Furthermore, our observed F_{ST} at rs702424 was not significantly different from mean F_{ST} of the 1069 SNPs from the 1M Illumina array based on the one-sample t -test ($t=22.8$, $df=1068$, $P>0.05$). By contrast, however, in a prior analysis of *TAS2R16* in a similar set of diverse African populations, we found that observed F_{ST} at polymorphic site 516 ($F_{ST}=0.076$) in the coding exon was an outlier (>98 th percentile) in the genome-wide distribution of F_{ST} values derived from the 1M Illumina data, consistent with a model of local adaptation at this locus.⁸

To increase power to accurately reconstruct haplotypes, we integrated genotype data at rs702424 with previously collected *TAS2R16* coding exon sequences for the same individuals ($N=165$), and we inferred nine distinct haplotypes. We then constructed a median-joining network (Figure 1), which showed that haplotype diversity grouped into two main clusters: (1) 'A-G'

haplotypes (defined by the ancestral A-allele at rs702424 and the ancestral G-allele at site 516 within *TAS2R16*) and (2) 'A-T' haplotypes (defined by the ancestral A-allele at rs702424 and the derived T-allele at site 516) and 'G-T' haplotypes (defined by the derived G-allele at rs702424 and the derived T-allele at site 516) (Table 1 and Figure 1). Our genealogy also revealed that 'G-T' haplotypes were present in divergent African populations with different diets and living in different geographic regions (Table 1; Figure 1), suggesting that these variants are quite old. However, given that the 'G-T' haplotypes are positioned furthest from the root of the tree (Figure 1), we argue that the derived G-allele at rs702424 is likely younger than the T-allele at site 516, which was previously estimated to be ~ 1 million years old.⁸ In addition, the reticulation among the 'A-T' and 'G-T' haplotypes in the 'high-sensitivity' clade in Figure 1 indicates that recombination has occurred between variation at rs702424 and the T-allele at site 516 in African populations.

To further explore signatures of selection based on long-range LD patterns, we applied the Extended Haplotype Homozygosity (EHH)¹⁴ statistic in African samples for which we had Illumina 1M Duo genotype and *TAS2R16* sequence data ($N=165$). Specifically, we integrated previously sequenced *TAS2R16* coding exon data with $\sim 60,000$ SNPs from the Illumina 1M Duo genotyped across the entire length of chromosome 7, and measured the decay of EHH from the core SNP (rs702424). However, we did not observe extensive EHH on chromosomes carrying the derived G-allele at rs702424 compared

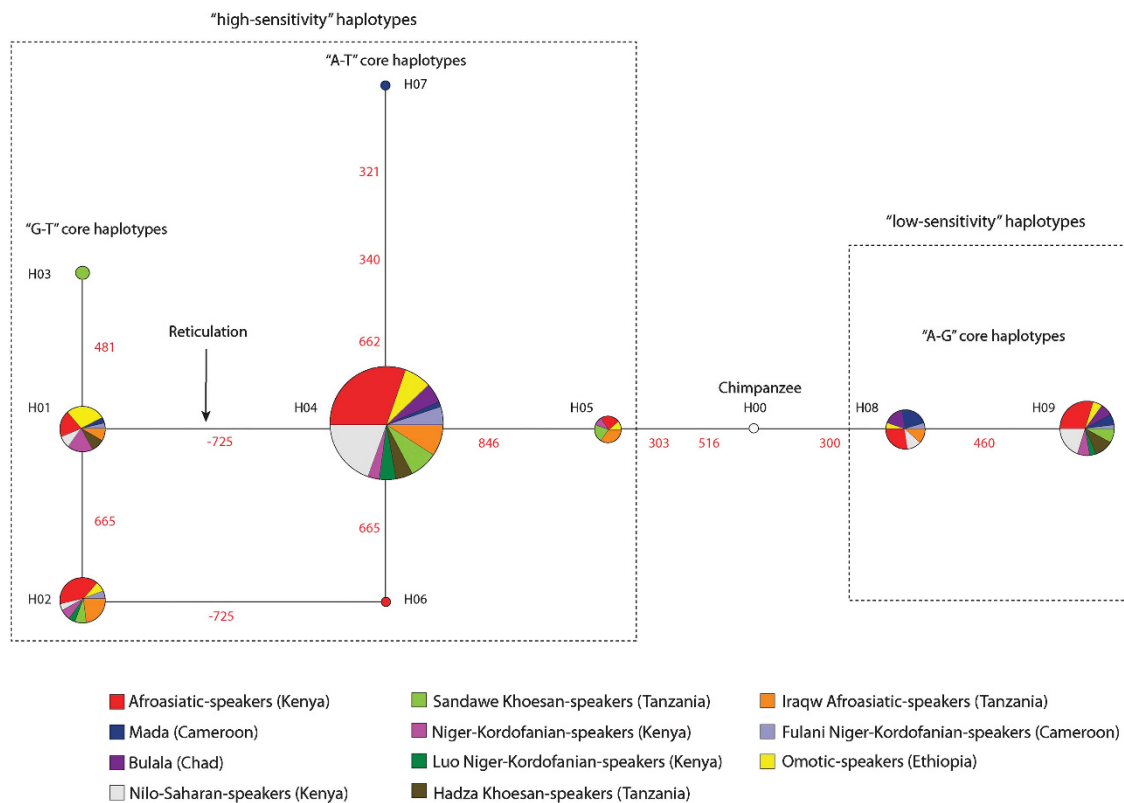


Figure 1 Genealogy of haplotype relationships in African populations. Here the circles represent haplotypes, and the size of the circles indicates the number of chromosomes in our samples with that particular haplotype. The colors within each circle also represent the proportion of that haplotype found in a particular population. The numbers between haplotypes indicate the nucleotide position where a mutation has occurred (-725 is the nucleotide position of rs702424 upstream from the transcription start site of *TAS2R16*). The dashed lines demarcate haplotypes previously associated with 'high' and 'low' salicin bitter taste sensitivity.^{8,10} In addition, we have indicated haplotypes defined by variation both at rs702424 in the *TAS2R16* promoter and at site 516 in the *TAS2R16* coding exon (for example, 'G-T' haplotypes correspond to the G-allele at rs702424 and the T-allele at site 516).

Table 1 Frequency of haplotypes defined by alleles in the *TAS2R16* promoter and coding regions

Population Clusters	'G-T' core haplotypes (H1–H3)			'A-T' core haplotypes (H4–H7)				'A-G' core haplotypes (H8–H9)	
	H1	H2	H3	H4	H5	H6	H7	H8	H9
Fulani (2N=16)	0.063	0.125	0.000	0.688	0.000	0.000	0.000	0.063	0.063
Mada (2N=10)	0.100	0.000	0.000	0.200	0.000	0.000	0.100	0.400	0.200
Bulala (2N=18)	0.000	0.000	0.000	0.556	0.000	0.000	0.000	0.167	0.167
Omotoc (Aari and Hamer) (2N=30)	0.267	0.067	0.000	0.533	0.033	0.000	0.000	0.033	0.067
Kenya_AA (2N=98)	0.051	0.102	0.000	0.653	0.020	0.010	0.000	0.051	0.112
Kenya_NS (2N=52)	0.038	0.019	0.000	0.769	0.000	0.000	0.000	0.038	0.135
Kenya_NK (2N=18)	0.278	0.111	0.000	0.389	0.056	0.000	0.000	0.000	0.167
Luo (2N=12)	0.000	0.083	0.000	0.833	0.000	0.000	0.000	0.000	0.083
Hadza (2N=16)	0.125	0.000	0.000	0.625	0.000	0.000	0.000	0.000	0.250
Sandawe (2N=26)	0.000	0.077	0.077	0.654	0.077	0.000	0.000	0.000	0.115
Iraqw (2N=32)	0.063	0.188	0.000	0.594	0.094	0.000	0.000	0.063	0.000

The first row in this table lists core haplotypes defined by variation at rs702424 in the *TAS2R16* promoter and at site 516 in the *TAS2R16* coding exon. The second row lists the inferred haplotypes (H1–H9) and the proportion of haplotypes present in distinct African populations is given below. Population clusters (which consist of groups of populations that share genetic and linguistic similarities) and the number of chromosomes (2N) analyzed are listed in the first column. NS, NK and AA refer to Nilo-Saharan, Niger-Kordofanian and Afroasiatic speakers. Genetically distinct populations within language families were classified as separate groups (for example, Fulani Niger-Kordofanian and Luo Nilo-Saharan speakers).

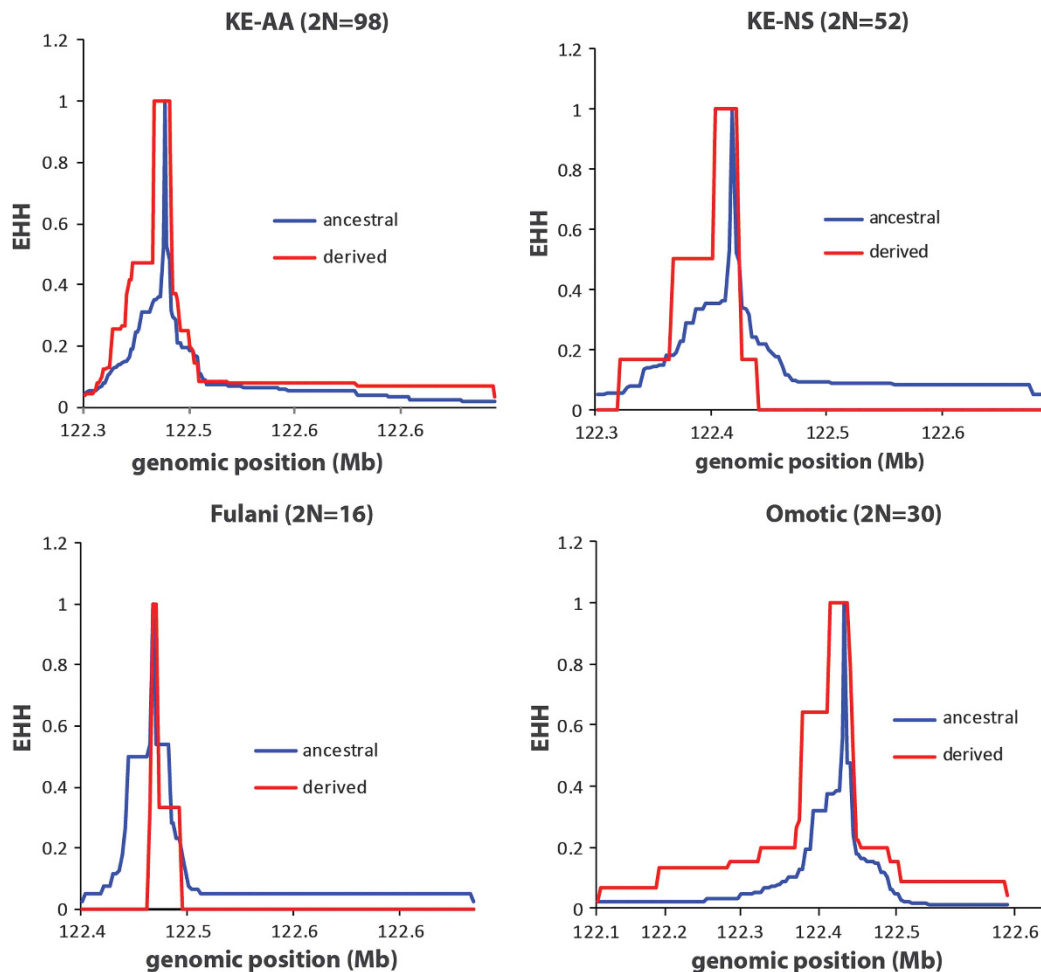


Figure 2 EHH plots for selected African populations. We applied the EHH statistic to identify signatures of recent positive selection based on long-range LD on chromosomes containing the rs702424 core SNP for individuals for whom we also had Illumina Human 1M-Duo genotype and *TAS2R16* sequence data (N=165 individuals). The decay of haplotype homozygosity on chromosomes was measured by calculating EHH for the core rs702424 SNP and the surrounding SNPs in the order of increasing distances on either side of the core (in Mb). EHH=0 means all extended haplotypes are different, whereas EHH=1 indicates that all extended haplotypes are the same. Here, the red line represents the decay of homozygosity on chromosomes carrying the derived G-allele at rs702424, whereas the blue line signifies the decay of homozygosity on chromosomes with the ancestral A-allele at the core site. KE-AA and KE-NS are abbreviations for Kenyan Afroasiatic speakers and Kenyan Nilo-Saharan speakers, respectively. Omotoc speakers and the Fulani are from Ethiopia and Cameroon, respectively. The number of chromosomes (2N) analyzed for each population is given in parentheses.

with chromosomes carrying the ancestral A-allele (Figure 2; Supplementary Figure 1), contrary to the pattern we would expect for a recent selective sweep of variation.

Levels of salicin (a known ligand of the TAS2R16 receptor) sensitivity were measured using a modification of the classic threshold recognition test¹⁵ in a subset of 135 individuals from East Africa. We examined the functional effect of common variation at rs702424 by comparing mean salicin score (transformed to a normal distribution) for three genotypic classes: (1) individuals homozygous for the derived G-allele at rs702424 and with at least one T-allele at site 516; (2) individuals heterozygous for variation at rs702424 and with at least one T-allele at site 516; and (3) individuals homozygous for the ancestral A-allele at rs702424 and with at least one T-allele at site 516 using a GLM univariate analysis of variance adjusting for age and sex. We observed a lower mean salicin score for individuals homozygous or heterozygous for the G-allele at rs702424 compared with individuals homozygous for the A-allele at the same site. However, the difference in mean score between these genotypic classes was not statistically significant ($P > 0.05$).

Although previous analyses detected signatures of positive selection at site 516 in the coding exon of *TAS2R16* in Africa,^{8,10} we did not find evidence of selection at rs702424 in the *TAS2R16* promoter based on EHH and F_{ST} statistics in a similar set of African populations in the present study. However, a study inferred recent selection for the derived mutations at rs702424 and site 516 primarily in Eurasian populations.⁹ Given the widespread global distribution of the derived G-allele at rs702424,⁹ we suggest that this mutation arose on haplotypes with the derived T-allele at site 516 before modern humans migrated from Africa ~40 000–80 000 years ago. If, indeed, selection occurred only on haplotypes carrying the derived alleles at both rs702424 and site 516 in non-Africans as previously proposed,⁹ selection at site rs702424 must have occurred after the expansion of modern humans out of Africa. Alternatively, the long-range LD structure present in non-Africans may have resulted from the population bottleneck that accompanied the geographic expansion from Africa.^{9,16} Nonetheless, it is clear that the G-allele at rs702424 has not been a target of positive selection in African populations unlike the derived T-allele at site 516. In addition, we did not observe a significant effect of alleles at rs702424 on salicin bitter taste recognition, supporting the recent finding that nonsynonymous variation at site 516 mainly influences salicin phenotypic variance. Although variability at rs702424 did not exhibit signals of selection in Africa and was not correlated with salicin bitter taste, continued research on *TAS2R16* and other *TAS2R* loci (including regulatory variation), together with phenotype data, will be important for understanding both the evolution and function of these understudied receptor genes.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Fathya Abdo, Eva Aluvalla, Godfrey Lema, Birhanu Mekauntintie, Alemayehu Moges, Hussein Musa, Lilian Nyindodo and Solomon Taye who helped with the field work, and the various institutions in West Central, Central and East Africa for their generous support. We thank Dongbo Hu and Wen-Ya Ko for useful discussions. We also offer our deepest appreciation to the many Africans who generously donated their DNA and time so that we can learn more about African genetic diversity. This work was supported by the US National Science Foundation (grant numbers BCS-0552486, BCS-0827436), US National Institutes of Health (grant numbers R01GM076637, 5DP1ES022577 05) to SAT and US National Institutes of Health (grant number R01 DC02995) to PASB.

- 1 Bufo, B., Hofmann, T., Krautwurst, D., Raguse, J. D. & Meyerhof, W. The human TAS2R16 receptor mediates bitter taste in response to beta-glucopyranosides. *Nat. Genet.* **32**, 397–401 (2002).
- 2 Greene, T. A., Alarcon, S., Thomas, A., Berdougou, E., Doranz, B. J., Breslin, P. A. *et al.* Probenecid inhibits the human bitter taste receptor TAS2R16 and suppresses bitter perception of salicin. *PLoS ONE* **6**, e20123 (2011).
- 3 Hayes, J. E., Wallace, M. R., Knopik, V. S., Herbstman, D. M., Bartoshuk, L. M. & Duffy, V. B. Allelic variation in TAS2R bitter receptor genes associates with variation in sensations from and ingestive behaviors toward common bitter beverages in adults. *Chem. Senses* **36**, 311–319 (2011).
- 4 Lee, R. J., Xiong, G., Kofonow, J. M., Chen, B., Lysenko, A., Jiang, P. *et al.* T2R38 taste receptor polymorphisms underlie susceptibility to upper respiratory infection. *J. Clin. Invest.* **122**, 4145–4159 (2012).
- 5 Prince, A. The bitter taste of infection. *J. Clin. Invest.* **122**, 3847–3849 (2012).
- 6 Shah, A. S., Ben-Shahar, Y., Moninger, T. O., Kline, J. N. & Welsh, M. J. Motile cilia of human airway epithelia are chemosensory. *Science* **325**, 1131–1134 (2009).
- 7 Campbell, M. C., Ranciaro, A., Froment, A., Hirbo, J., Omar, S., Bodo, J. M. *et al.* Evolution of functionally diverse alleles associated with PTC bitter taste sensitivity in Africa. *Mol. Biol. Evol.* **29**, 1141–1153 (2012).
- 8 Campbell, M. C., Ranciaro, A., Zinshteyn, D., Rawlings-Goss, R., Hirbo, J., Thompson, S. *et al.* Origin and differential selection of allelic variation at TAS2R16 associated with salicin bitter taste sensitivity in Africa. *Mol. Biol. Evol.* **31**, 288–302 (2014).
- 9 Li, H., Pakstis, A. J., Kidd, J. R. & Kidd, K. K. Selection on the human bitter taste gene, TAS2R16, in Eurasian populations. *Hum. Biol.* **83**, 363–377 (2011).
- 10 Soranzo, N., Bufo, B., Sabeti, P. C., Wilson, J. F., Weale, M. E., Marguerie, R. *et al.* Positive selection on a high-sensitivity allele of the human bitter-taste receptor TAS2R16. *Curr. Biol.* **15**, 1257–1265 (2005).
- 11 Wooding, S., Kim, U. K., Bamshad, M. J., Larsen, J., Jorde, L. B. & Drayna, D. Natural selection and molecular evolution in PTC, a bitter-taste receptor gene. *Am. J. Hum. Genet.* **74**, 637–646 (2004).
- 12 Tishkoff, S. A., Reed, F. A., Friedlaender, F. R., Ehret, C., Ranciaro, A., Froment, A. *et al.* The genetic structure and history of Africans and African Americans. *Science* **324**, 1035–1044 (2009).
- 13 Fst 2.9.3.2. [updated 2005; cited 2011 Jun 22]. Available at <http://www2.unil.ch/popgen/softwares/fst2.htm> (2002).
- 14 Sabeti, P. C., Walsh, E., Schaffner, S. F., Varrilly, P., Fry, B., Hutcheson, H. B. *et al.* The case for selection at *CCR5-Delta32*. *PLoS Biol.* **3**, e378 (2005).
- 15 Harris, H. & Kalmus, H. The distribution of taste thresholds for phenylthiourea of 384 sib pairs. *Ann. Eugen.* **16**, 226–230 (1951).
- 16 Schmegegn, C., Hoegel, J., Vogel, W. & Assum, G. Genetic variability in a genomic region with long-range linkage disequilibrium reveals traces of a bottleneck in the history of the European population. *Hum. Genet.* **118**, 276–286 (2005).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)